May 2003 • NREL/CP-500-33956

Conference Paper

Statistical Wind Power Forecasting Models: Results for U.S. Wind Farms

Preprint

M. Milligan, M. Schwartz, Y. Wan

To be presented at WINDPOWER 2003
Austin, Texas
May 18-21, 2003

National Renewable Energy Laboratory

1617 Cole Boulevard
Golden, Colorado 80401-3393

NREL is a U.S. Department of Energy Laboratory
Operated by Midwest Research Institute • Battelle • Bechtel

Contract No. DE-AC36-99-GO10337

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: reports@adonis.osti.gov

Available for sale to the public, in paper, from:
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

Printed on paper containing at least 50% wastepaper, including 20% postconsumer waste

STATISTICAL WIND POWER FORECASTING MODELS: RESULTS FOR U.S. WIND FARMS

Michael Milligan, Consultant
Marc Schwartz
Yih-Huei Wan

National Wind Technology Center
National Renewable Energy Laboratory
1617 Cole Blvd.
Golden, CO 80401
303-384-6927
303-384-6901 fax
mailto:Michael_Milligan@nrel.gov

ABSTRACT

Electricity markets in the United States are evolving. Accurate wind power forecasts are beneficial for wind plant operators, utility operators, and utility customers. An accurate forecast makes it possible for grid operators to schedule the economically efficient generation to meet the demand of electrical customers. In the evolving markets some form of auction is held for various forward markets, such as hour ahead or day ahead. In California, the Independent System Operator (ISO) filed a tariff for intermittent generators that uses independently developed wind power forecasts to help with resource selection in the hour-ahead market. The Federal Energy Regulatory Commission (FERC) has recommended inclusion of this tariff in its Standard Market Design. This paper develops several statistical forecasting models that can be useful in hour-ahead markets that have a similar tariff. Although longer-term forecasting relies on numerical weather models, the statistical models used here focus on the short-term forecasts that can be useful in the hour-ahead markets. The purpose of the paper is not to develop forecasting models that can compete with commercially available models. Instead, we investigate the extent to which time-series analysis can improve on simplistic persistence forecasts. This project applied a class of models known as autoregressive moving average (ARMA) models to both wind speed and wind power output. The ARMA approach was selected because it is a powerful, well-known time-series technique and has been used by the California Independent System Operator in some of its forecasting work. Results are presented for operating wind farms in Iowa and Minnesota, and indicate that a significant improvement over persistence models is sometimes possible.

INTRODUCTION

In some electricity markets wind is becoming a significant source of energy. As the use of wind power plants continues to grow, the impact of wind on various aspects of power system operation receives greater scrutiny. Because wind is an intermittent power source, these operational impacts are unlike those of other power plants. This intermittent characteristic of wind power generation means that efficient power system operation will depend in part on the ability to forecast available wind power.

A wind farm operator would want to perform some type of cost-benefit analysis before embarking on, or participating in, a forecasting project. Central to this cost-benefit study would be any relevant market constraints or other considerations, such as the monthly netting of imbalances that is part of the California ISO's wind tariff (this applies only to wind facilities that participate in the ISO's program and pay a forecasting fee).

This paper examines the use of a standard class of statistical time-series models to predict wind power output up to six hours in advance. The purpose is not to develop models to compete with commercial forecasting models, such as those described in [1] or [2]. Rather, our goal to investigate the feasibility of relatively inexpensive statistical forecasting models that do not require any data beyond historical wind power generation data. This may limit the ability and the usefulness of this type of forecasting model, but for small wind farms that are unable to participate in formal forecasting projects, it might be desirable to use a statistical model that can be developed and used at lower cost.

For this project we used data collected from operating wind power plants in Iowa and Minnesota. We found some differences in model performance at the two sites, and we would expect that sites in different climatic regions would exhibit even more variations. The data used for this project have been described in[3].

STATISTICAL MODELING FRAMEWORK

Although many time-series methods could be applied to this problem, a general class of models known as autoregressive integrated moving average (ARIMA) models is applied in this paper. Similar models have been applied by [4], [5], and [6]. ARIMA models have up to three components: autoregressive, integrated, and moving average. When the second component provides no significant explanatory power in the model, it is dropped. We found the integration term to be unimportant in this analysis, so it was dropped. The resulting model is known as ARMA, and can be characterized as

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \sum_{k=0}^{q} b_k e_{t-k}$$

The equation states that a realization of the time-series X at time t depends on a linear combination of past observations of X plus a moving average of series e, which is a white noise process characterized by zero mean and variance sigma. The time series X is known as an ARMA(p,q) process, where p is the order of the autoregressive process of X on itself, and q is the order of the moving-average error term.

There are well-known methods for applying ARMA models in practice. Interested readers can consult [7]for details. The first step of ARMA model development is to determine the order of the AR and MA processes, p and q, respectively. This is known as the *model identification* phase, and it involves the analysis of the autocorrelogram and partial autocorrelogram. Model identification is not an exact science, so it may be difficult to eliminate all the alternative model specifications.

Once tentative values of p and q have been identified, the model coefficients a_j and b_k can be estimated. Although the model is linear in X_{t-j} and e_{t-j}, nonlinear estimation methods are required because of the autoregressive nature of X. The use of a linear estimation technique would result in a biased estimate of the coefficients.

After estimation, the model can be checked with several diagnostics. It is not uncommon for a promising model identification to lead to a poorly performing model, so the diagnostic phase is important because it weeds out models that do not work well. We found many such models during this project. We chose to present results for similar models whenever the model-forecasting performance improved upon the persistence model.

METRICS

A well-known method of forecasting wind is the simplistic persistence method. This approach uses the past hour wind speed (or wind power) as the forecast for the next hour. As a forecasting technology, this method is not impressive, but it is nearly costless, and performs surprisingly well. Therefore, any forecast method first should be measured by the extent it can improve on persistence forecasts. That is the approach that we apply in this paper. The persistence forecast can offer a range of forecasting accuracy, depending on the wind regime and the number of periods to be forecast. In this project, we calculated the root mean square error (RMSE) of each forecast over the relevant time period to compare our methods with the persistence model. A lower RMSE implies that the forecast is more accurate, whereas a high RMSE value implies less accuracy.

We are interested in how well ARMA models can forecast more than 1 hour in advance. We applied our forecasting techniques up to 6 hours ahead, which is approximately the limitation of purely statistical forecasting methods. We also developed ARMA models to forecast in 10-minute intervals. This may not be useful in some power markets, but it does illustrate that there is potential for forecast periods of less than 1 hour. In all cases, we measured the model forecast performance by comparing RMSE for the relevant data.

HOURLY POWER FORECASTS

We evaluated the model performance on different data than were used to train (fit the parameters) the model in each of the forecasting cases we looked at. A number of different training periods were applied so that we could see how this would affect the forecasting performance.

We used actual wind power data from Lake Benton II (LB), Minnesota, as our primary source in this project. We also compared some of the LB forecasts with forecasts at Storm Lake, Iowa. Data from both of these sites is collected at the National Renewable Energy Laboratory and are described more fully in [3]. Figure 1 depicts the general location of these sites.

A number of decisions must be made on how to develop statistical hourly power forecasts in order to analyze their accuracy. One of the most important decisions is the selection of the training period for the forecasting algorithm. In some of our initial screening work using the LB data, we calculated separate time-series for 10 months, training the model for the first 2 weeks and calculated forecasts for the remainder of the month. The statistical evidence showed that there was a significant difference in the ARMA model specifications that were calculated in each of the training periods. This finding is not surprising because it provides evidence of statistically significant variations in wind generation patterns at different times of the year, and these differences are important in developing a forecasting model.

The first set of cases used LB wind power data from January-February 2001. The data from January were used to train the model, and hourly forecasts were developed for the month of February. Each forecast was for a 6-hour period ahead. The model results were compared with the persistence model, which was also applied to the same time periods. In many cases, the ARMA forecasting results were similar for different model specifications. We report results for forecast models that did improve over persistence.

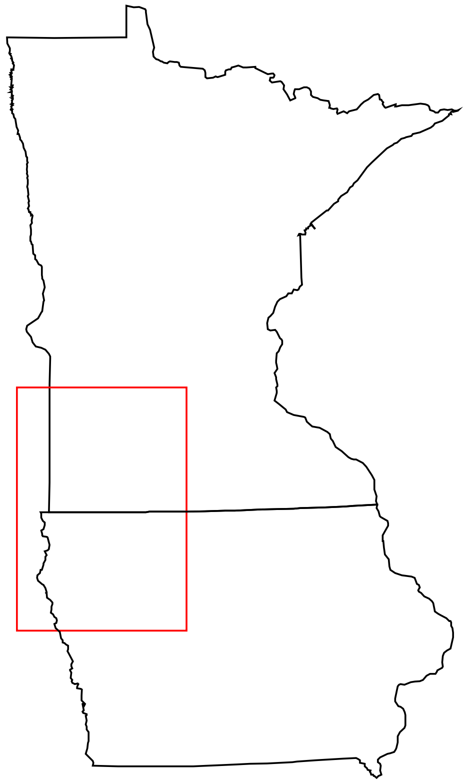

FIGURE 1. WIND FARM LOCATIONS.

Figure 2 illustrates the first forecasting case. The graph shows the ratio of the persistence RMSE to the ARMA RMSE. Higher ratios indicate better ARMA performance, and ratios less than 1 would indicate that persistence does a better job of forecasting than the ARMA model. The best model specification is clearly the ARMA(1,24), which improves on persistence by approximately 7% in the first hour, increasing to approximately 18% in the sixth hour.

Figures 3 and 4 illustrate the forecast degradation even of the best ARMA(1,24) model. As indicated in Figure 3, a 1-hour forecast is quite good, but the forecast performance is significantly worse for 2 hours ahead.

A similar set of forecasting models were applied to the same time period at the Storm Lake wind plant. The same model specification, the ARMA(1,24), did the best overall job. However, it is worth noting that the ARMA(1,2) model was best in the 1-hour forecast, whereas the ARMA(1,24) model was better from 2 hours out to 6 hours. This suggests the possibility of using an ensemble of ARMA models, depending on the forecast performance of different model specifications for different forecast horizons. The Storm Lake results appear in Figure 5.

The ARMA forecasts at both wind sites eroded significantly as the number of forecast periods increased, although the Storm Lake results were slightly better than the LB results. As measured by the RMSE, the forecast error for LB nearly tripled from the one-hour forecast to the 6-hour forecast, and the Storm Lake results were similar. A graph that shows the RMSE in kW as a function of the length of the forecast period appears in Figure 6.

We also investigated additional periods during 2001; and changed the lengths of the training and testing intervals. Figure 7 shows an example of LB for the month of April 2001. For this case, a 3-week training period was used, and the remaining hours of the month were used to test the forecast. For this period, it is apparent that the suite of tested models did improve on persistence, but not to the same degree as the January-February cases.

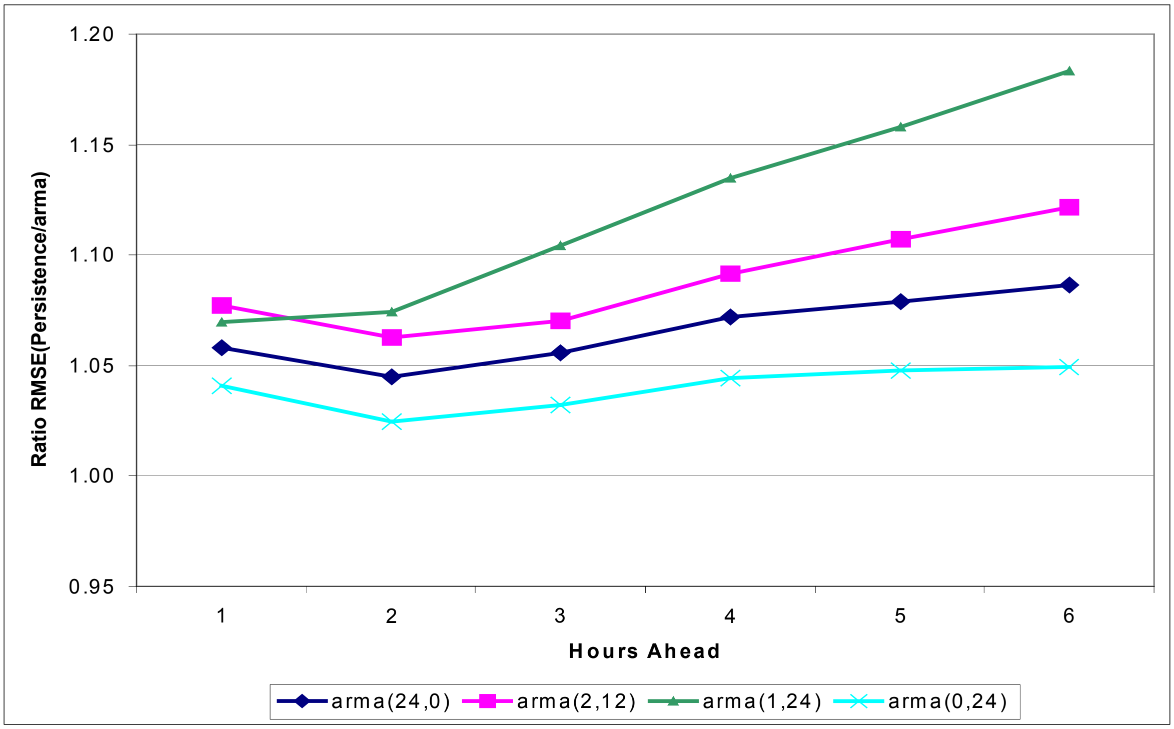

FIGURE 2. LAKE BENTON II kW FORECASTS: JANUARY/FEBRUARY 2001.

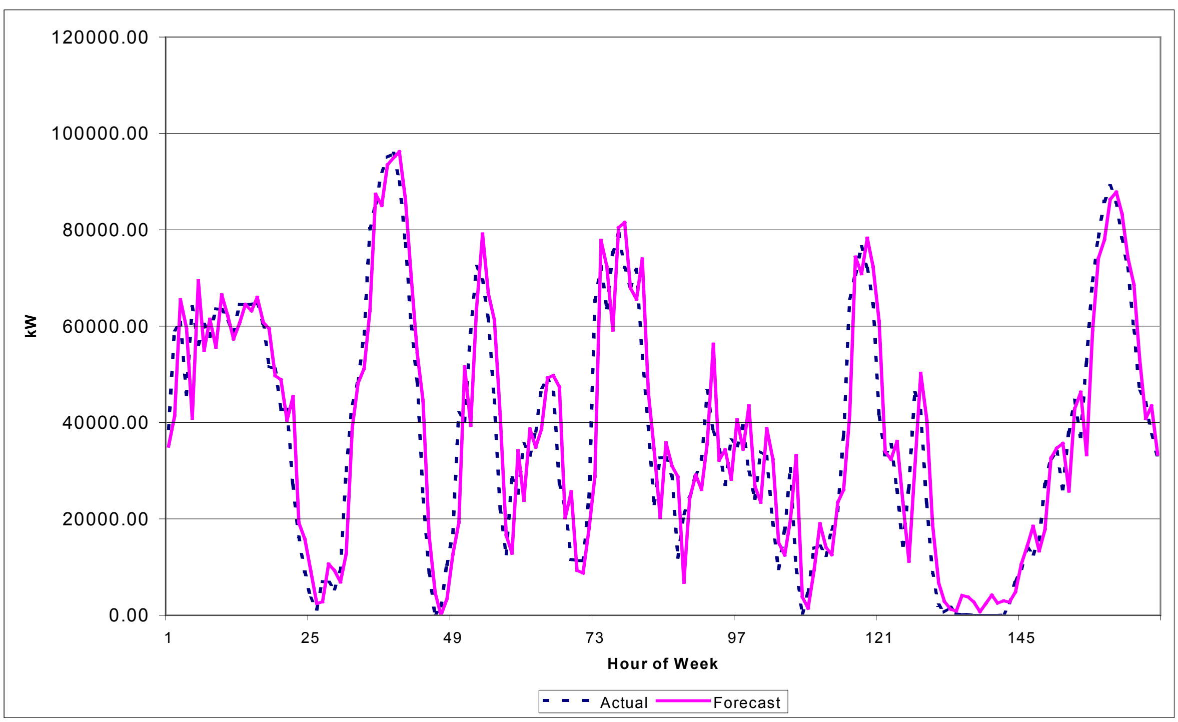

FIGURE 3. LAKE BENTON II kW 1-HOUR FORECASTS vs. ACTUAL: JANUARY/FEBRUARY 2001. ARMA(1,24).

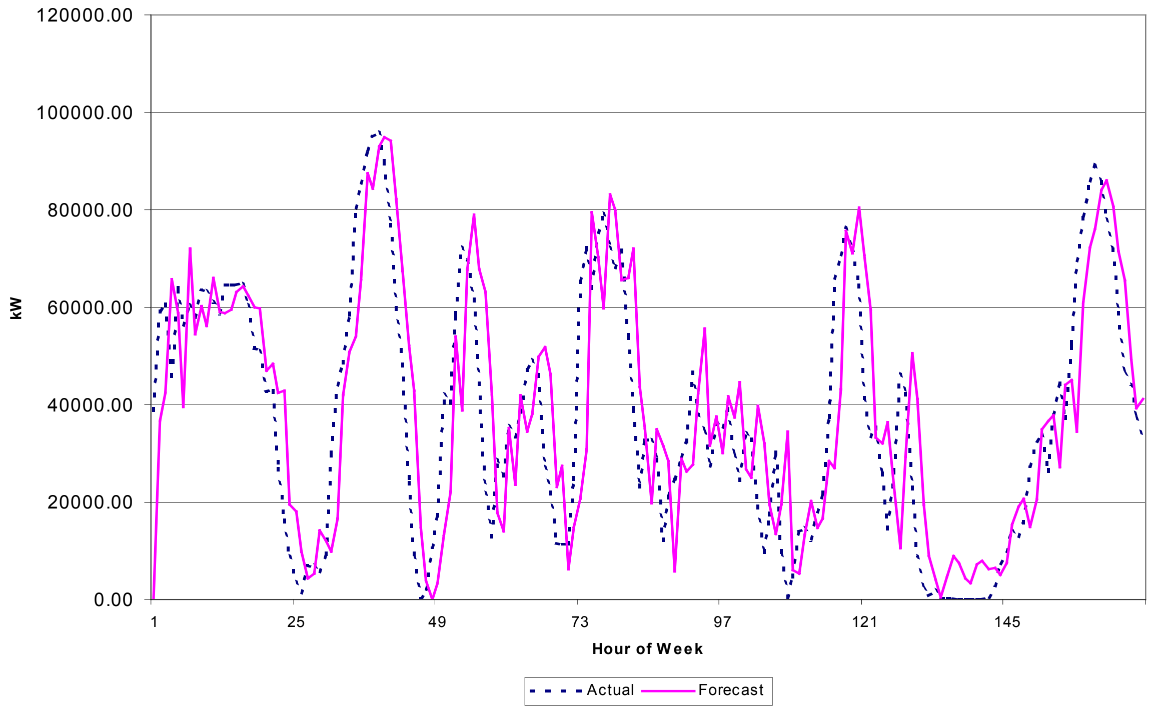

FIGURE 4. LAKE BENTON II kW 2-HOUR FORECASTS vs. ACTUAL: JANUARY/FEBRUARY 2001.

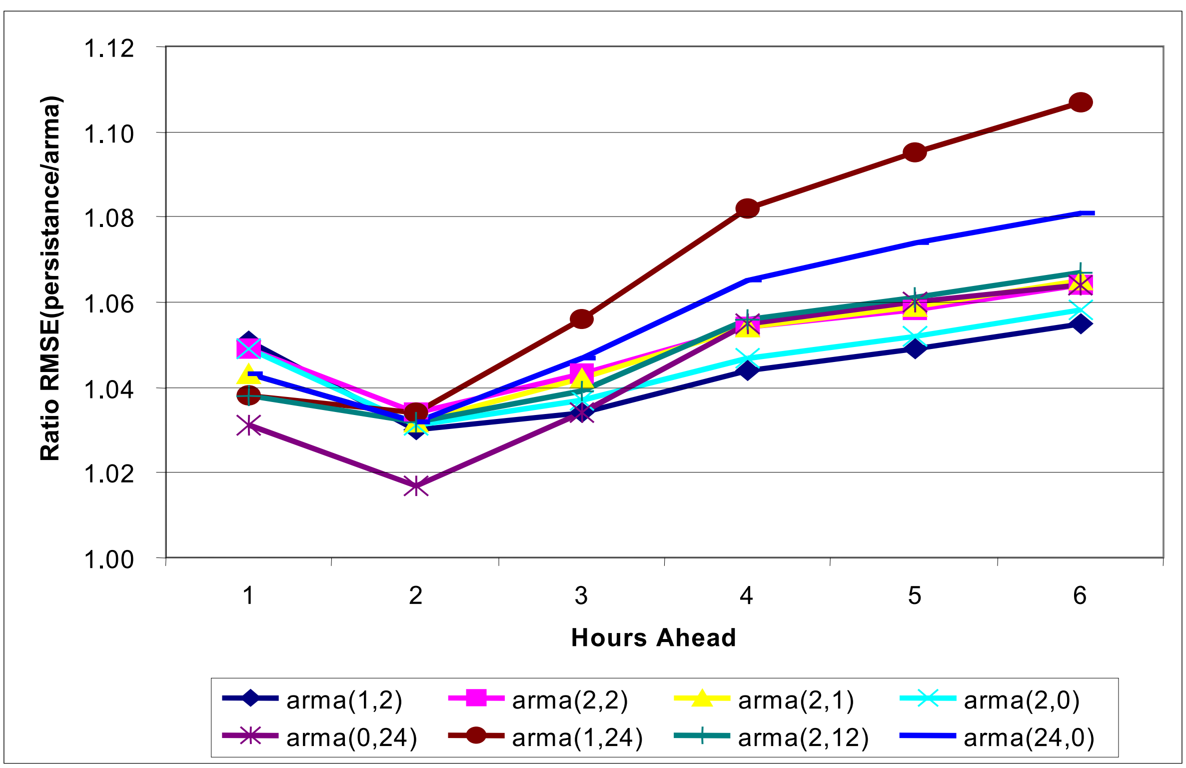

FIGURE 5. STORM LAKE kW FORECASTS: JANUARY/FEBRUARY 2001.

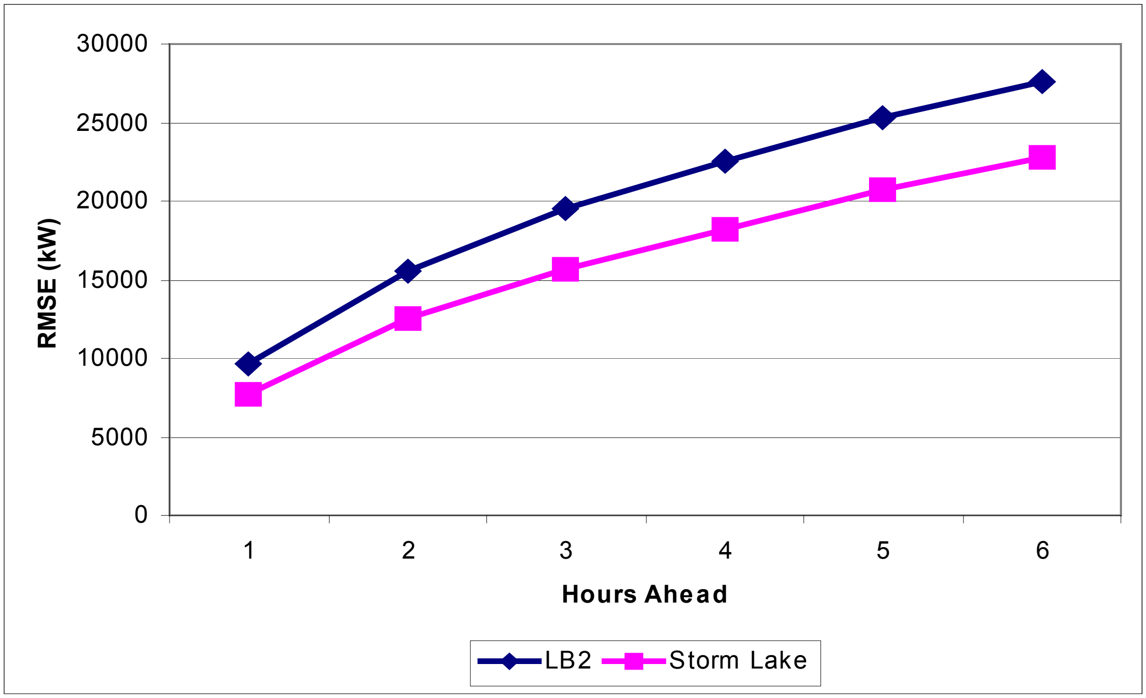

FIGURE 6. ARMA(1,24) kW FORECAST DEGRADATION: JANUARY/FEBRUARY 2001.

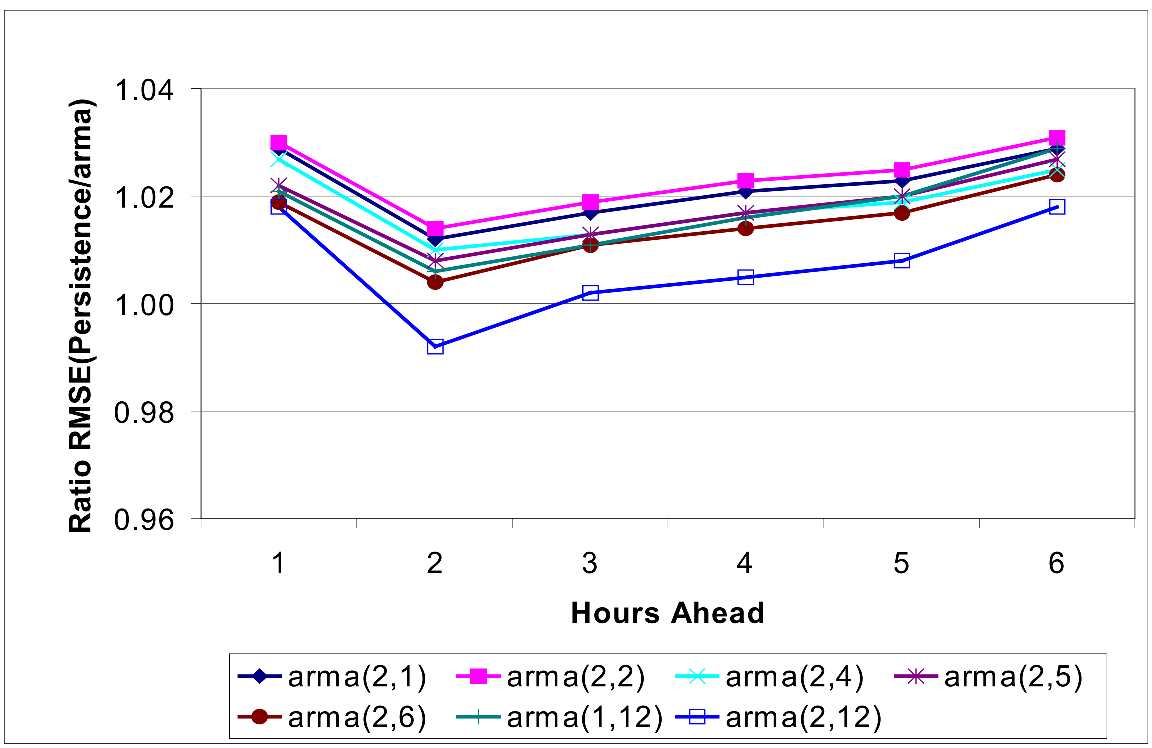

FIGURE 7. LAKE BENTON II kW FORECASTS: APRIL 2001.

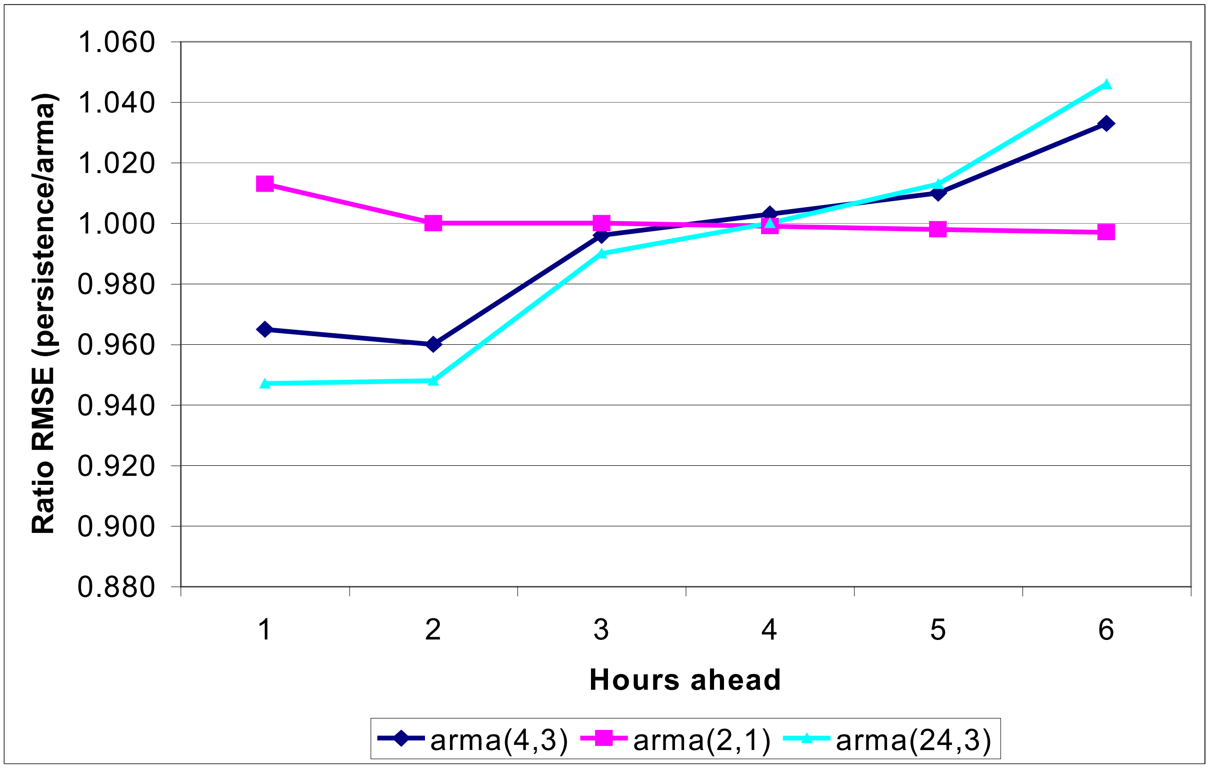

FIGURE 8. LAKE BENTON II WIND SPEED FORECASTS: MARCH/APRIL 2001.

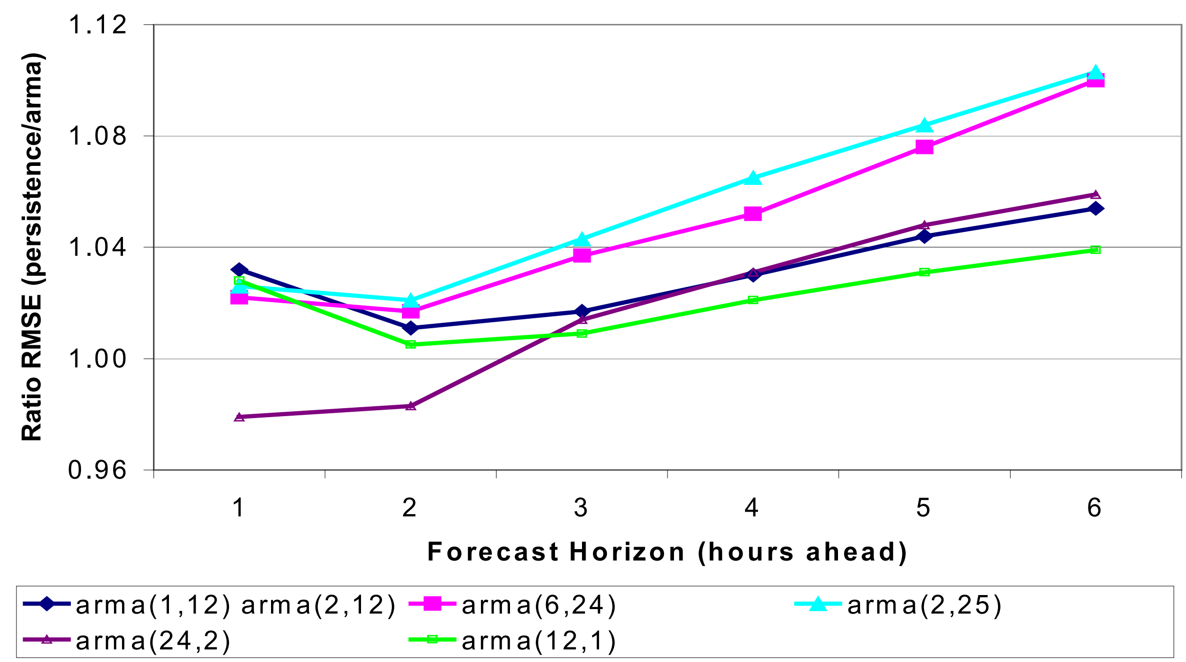

FIGURE 9. LAKE BENTON II WIND SPEED FORECASTS: APRIL/MAY 2001.

Similar differences in model-forecasting ability at different times were found when forecasting wind speed directly. For LB we used March 2001 as a training period and tested the forecast for the following month. During the ARMA model identification phase the order of the ARMA process was somewhat ambiguous. Approximately 175 different ARMA specifications were tested, and none of them offered a particularly impressive forecast. These results appear in Figure 8.

April 2001 was used for the training period, and May was used to test the forecasts for the next case. The forecasting performance in the first couple of hours improved on persistence by about 3% (or less) based on the RMSE, increasing to about 11% for the 6-hour forecast. Results from this suite of forecast models can be seen in Figure 9. These results are improved when compared to the March/April 2001 results.

DYNAMIC TRAINING PERIODS

Because of the apparent forecast model sensitivity to time of year, we also applied a dynamic training period to some of the LB cases. A more complete study than ours might look at a wide range of dynamic training periods, ranging from perhaps 1 week to several weeks. For the first dynamic training period, we used 744 hours (January 2001) to train the model, and applied the forecast to February. The approach uses the previous 744 hours to fit the model parameters, after first identifying the order of the ARMA process. A 6-hour forecast is generated. Then the training period shifts forward by 1 hour and another 6-hour forecast is generated. The process continues for the full month of February. Figure 10 illustrates the results of this process. For the 1-hour forecast, the range of improvement over persistence goes from 7% for the 1-hour forecast, up to 15% for the 6-hour forecast. Note that the same ARMA model does not consistently perform better than the others in this suite.

We chose the March-April time period for LB and applied different dynamic training periods to investigate this issue in more detail. This period was chosen because it appears to be the most difficult to predict, and we would like to see how the dynamic training period selection might help improve the forecasts. We show three cases, corresponding to dynamic training periods of 2 weeks, 3 weeks, and the full month of March. Aside from the statistical performance of the models, different training period lengths have different tradeoffs between the information provided to the model and the constraints that are imposed by the model parameters. For example, a single ARMA model that is fit over a period such as 1 year will potentially have the ability to use a great deal of information that is embedded in the wind signal. Because different physical mechanisms can affect the level of wind resource at different times of the year, the variability of the wind can also have different statistical properties during these periods. This implies that either a very large number of model parameters may be needed to fully describe the process, or the parameters themselves could be a function of time. Partitioning the year into a number of distinct time periods allows the model parameters to be re-fit to account for different climatological properties during different times of the year. Choosing too short a period for model training could leave out some important information that would help the model's forecasting accuracy. The ideal training period would pick up the important drivers and patterns for different times of the year. Parameters based on one set of climatic drivers should not be imposed on other time periods if it is known that another set of climatic drivers affects the wind resource. When model training and application spans a distinct seasonal boundary for example, we expect the forecasting performance to suffer.

Figure 11 shows the results of a dynamic 2-week training period, beginning in March 2001 and extending through the month of April. The ARMA(2,1) model does the best job of forecasting, but it only offers a 2%-5% improvement over persistence. Figure 12 shows a similar suite of ARMA models and is based on a 3-week dynamic training period. The forecast results are generally worse than the 2-week period. Figure 13 shows a 744-hour dynamic training period.

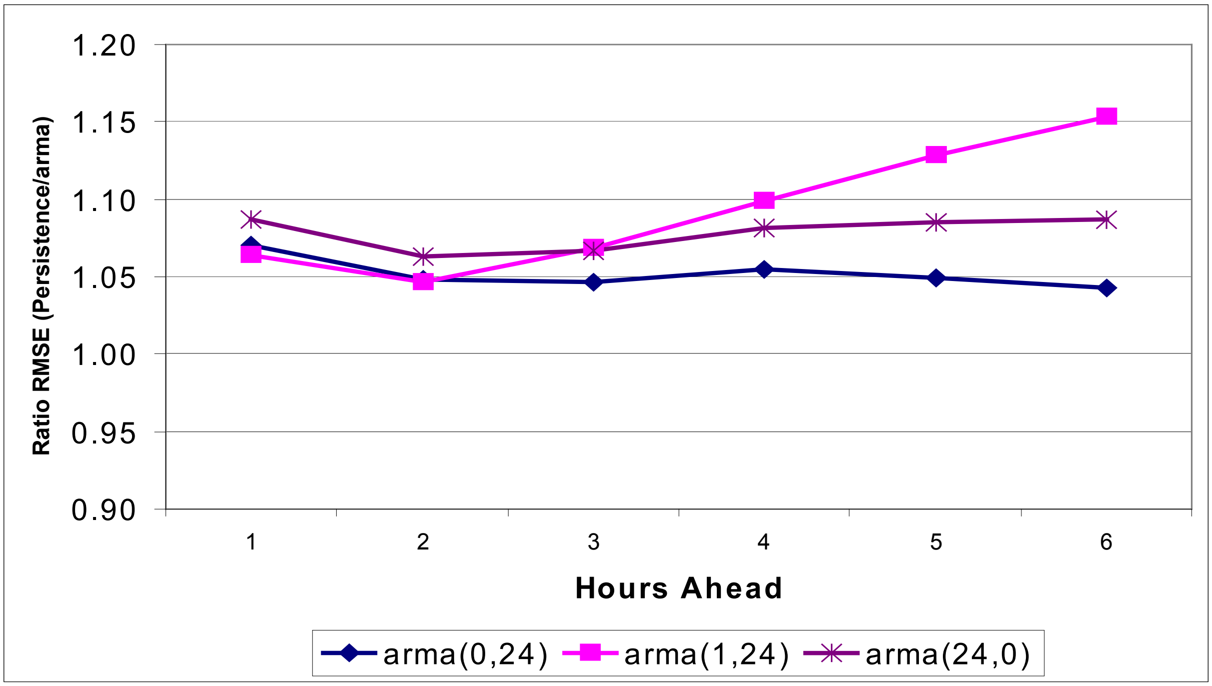

FIGURE 10. LAKE BENTON II kW FORECASTS WITH 744-HOUR DYNAMIC TRAINING: JANUARY/FEBRUARY 2001.

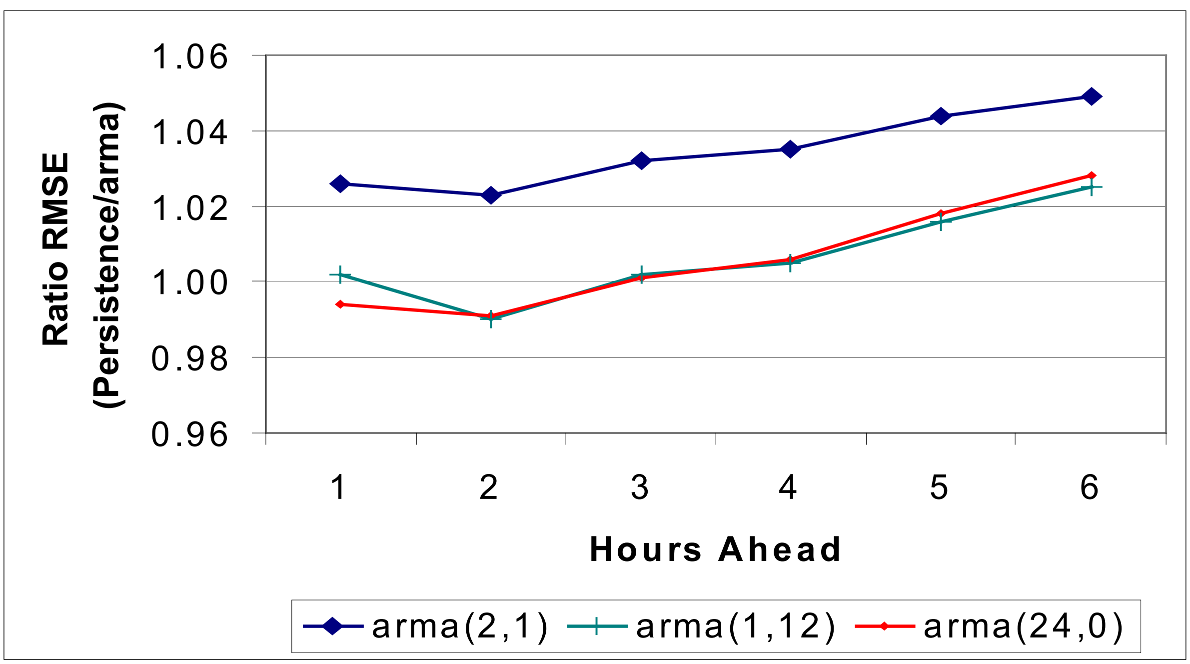

FIGURE 11. LAKE BENTON II kW FORECASTS WITH 2-WEEK DYNAMIC TRAINING: MARCH/APRIL 2001.

The results are moderately better than for the 3-week period, but the improvement over persistence is not significant.

The results from using the dynamic training periods are complex and not open to easy interpretation. Further investigation into this issue is certainly warranted.

10-MINUTE FORECASTS

Finally, we did some testing of a 10-minute-ahead forecasting model. The usefulness of this type of wind forecast model would vary depending on the market structure of the electric power industry. For example, in markets such as California that perform re-dispatching on a 10-minute interval, a 10-minute wind power forecast might be useful to determine whether to accept increment or decrement bids from generators. If the 10-minute wind forecast indicates less wind energy than previously scheduled, the power system operator might want to increase generation from other units to make up the shortfall; and would thus accept a previously submitted bid to increase conventional generator output. We would like to determine whether it is feasible to predict 10-minute wind power output with an ARMA model. We applied a suite of ARMA models to Lake Benton II data from April. We used a 3-week training period and applied the model to a 1-week period, generating 12 successive 10-minutes forecasts. Performance of several of these models was reasonably good. Figure 14 shows an ARMA(0,36) model applied to the 10-minute forecasting problem. Although several other models provided an advantage over the persistence model, their performance deteriorated rapidly after the first 10-minute period. The ARMA(0,36) model achieved a 16% RMSE improvement over persistence for one period ahead, falling to about 8% three periods ahead, and about 7.5% four periods ahead. At eight periods ahead, the ARMA(0.36) model deteriorates to the same performance level as the persistence model. Figure 15 shows a series of concatenated one-period (10-minute) forecasts from this model.

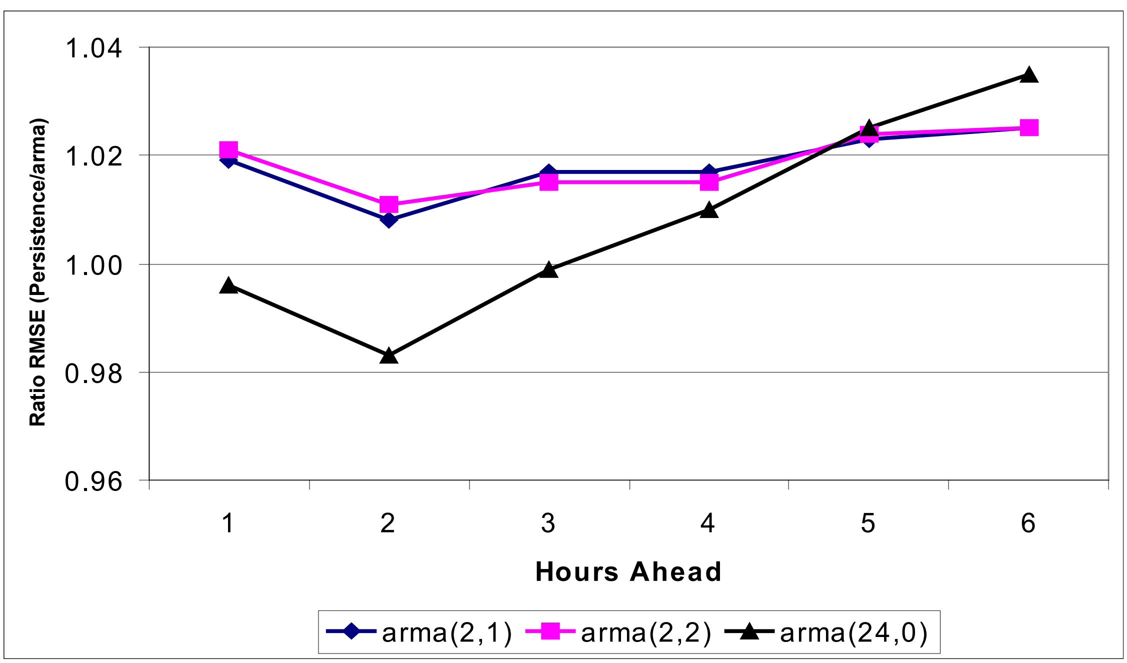

FIGURE 12. LAKE BENTON II kW FORECASTS WITH 3-WEEK DYNAMIC TRAINING: MARCH/APRIL 2001.

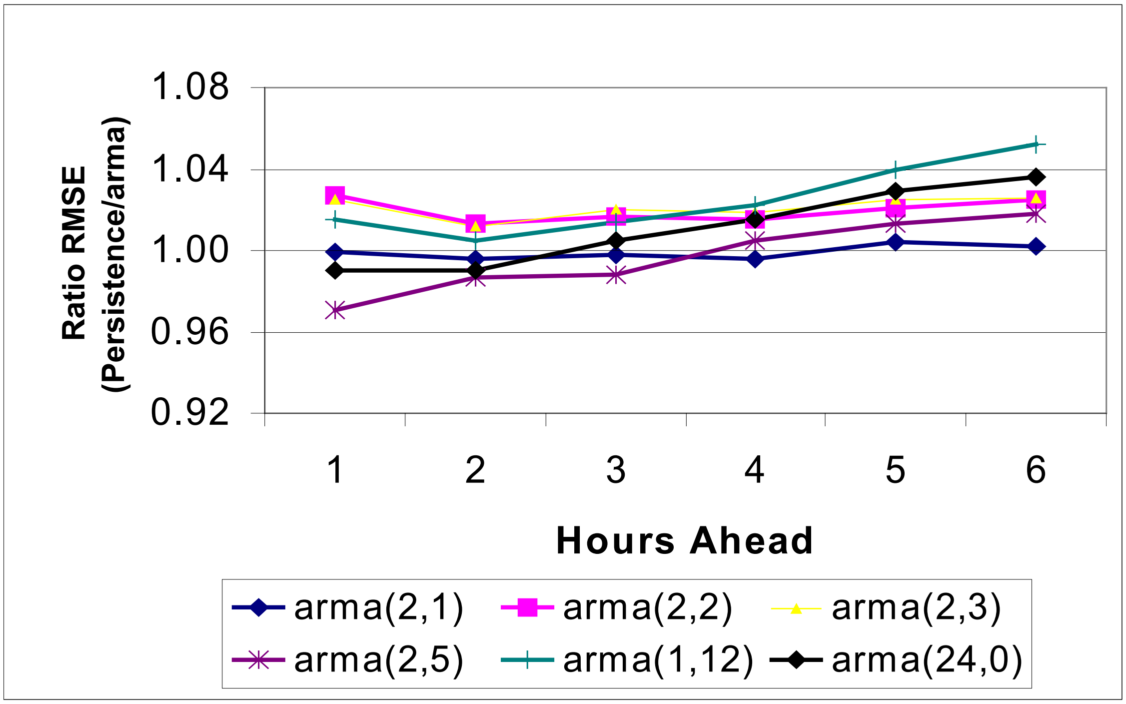

FIGURE 13. LAKE BENTON II kW FORECASTS WITH 744-HOUR DYNAMIC TRAINING: MARCH/APRIL 2001.

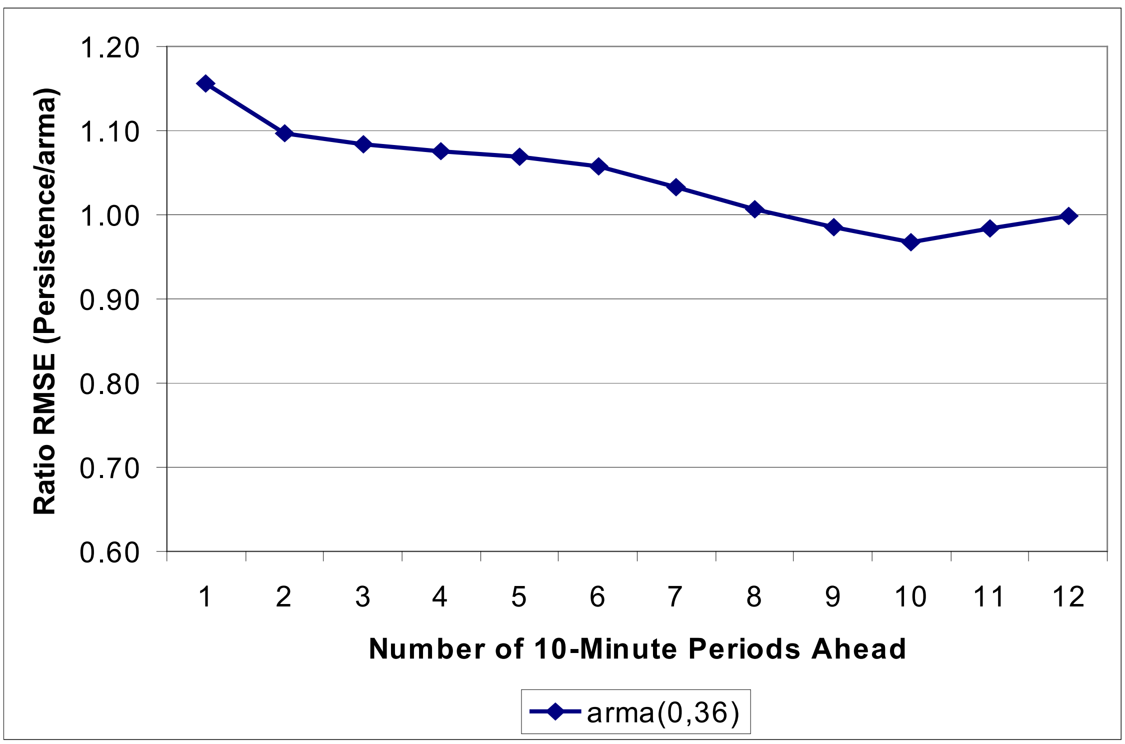

FIGURE 14. LAKE BENTON II 10-MINUTE kW FORECASTS APRIL 2001.

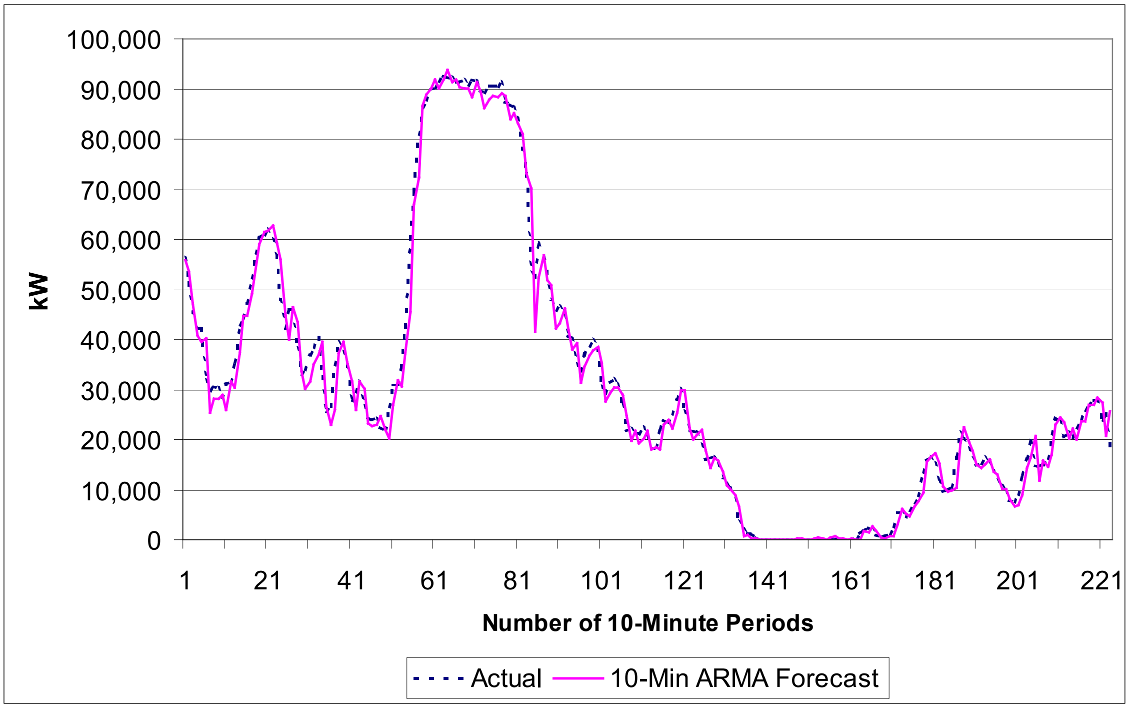

FIGURE 15. LAKE BENTON II 10-MINUTE kW FORECASTS vs. ACTUAL: APRIL 2001, ARMA(0,36).

CONCLUSIONS

There is a clear difference in the ability of ARMA forecast models applied to different time periods. The Lake Benton II wind farm was much easier to predict in January-February 2001, than it was in March-April 2001. Some of the models developed in this paper offer a significant improvement over the persistence model. Some ARMA models have difficulty reducing the forecast RMSE significantly as compared to persistence. In some cases, the model that does the best job forecasting 1-2 hours out is eclipsed by another model for longer forecast horizons. This raises the possibility of simple ensemble forecasts that can offer better forecasting over the horizon than a single model specification.

In several cases, we found many alternative ARMA models that did a good job forecasting over the testing time frame. We also found it difficult to determine the proper identification for several models, and that led to the evaluation of alternative specifications. It is apparent that a one-size-fits-all approach will not work based on the analysis of the role of the training period because of model performance sensitivity to this feature. The dynamic-training model applied to LB for January-February did not quite achieve the forecast accuracy as the fixed training period version of the model, suggesting that important information may have been excluded as the training window advanced forward.

Finally, we found that it is possible to forecast wind power 10 minutes ahead, and in 10-minute periods. The ARMA(0.36) model for LB in April 2001 beat persistence for up to eight 10-minute periods ahead. Although 10-minute forecasts may not be useful to all wind farm or grid operators, it might be useful in some markets.

REFERENCES

1. Bailey, B., M. Brower, J. Zack. *Short-term Wind Forecasting: Development and Application of a Mesoscale Model*. *European Wind Energy Conference*. 1999. Nice, France. p. 1062-1065.
2. Landberg, L. *Predicting the Power Output From Wind Farms*. *European Wind Energy Conference*. 1997. Dublin Castle, Ireland: EWEA. p. 747-750.
3. Wan, Y., Bucaneg, D. *Short-Term Power Fluctuations of Large Wind Power Plants*. *21st AMSE Wind Energy Symposium*. 2002. Reno, Nevada: NREL/CP-500-30747 (preprint).
4. Makarov, Y., D. Hawkins, E. Leuze, J. Vidov. *California ISO Wind Generation Forecasting Service Design and Experience*. *Windpower 2002*. 2002. Portland, OR: AWEA. CD-ROM.
5. Nielsen, T.H.M. *Experiences With Statistical Methods for Wind Power Prediction*. in *European Wind Energy Conference*. 1999. Nice, France: EWEA. p. 1066-1069.
6. Kariniotakis, G., N.E. Nogaret, G. Stavrakakis. *Advanced Short-Term Forecasting of Wind Power Production*. in *European Wind Energy Conference*. 1997. Dublin Castle, Ireland: EWEA. p. 751-754.
7. Box, G.E.P., and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*. Revised ed. 1976, San Francisco: Holden-Day.

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2003	Conference Paper

4. TITLE AND SUBTITLE

Statistical Wind Power Forecasting Models: Results for U.S. Wind Farms; Preprint

5. FUNDING NUMBERS

WER3 3610

6. AUTHOR(S)

M. Milligan, M. Schwartz, Y. Wan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

National Renewable Energy Laboratory
1617 Cole Blvd.
Golden, CO 80401-3393

8. PERFORMING ORGANIZATION REPORT NUMBER

NREL/CP-500-33956

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Electricity markets in the United States are evolving. Accurate wind power forecasts are beneficial for wind plant operators, utility operators, and utility customers. An accurate forecast makes it possible for grid operators to schedule the economically efficient generation to meet the demand of electrical customers. In the evolving markets, some form of auction is held for various forward markets, such as hour ahead or day ahead. This paper develops several statistical forecasting models that can be useful in hour-ahead markets that have a similar tariff. Although longer-term forecasting relies on numerical weather models, the statistical models used here focus on the short-term forecasts that can be useful in the hour-ahead markets. We investigate the extent to which time-series analysis can improve on simplistic persistence forecasts. This project applied a class of models known as autoregressive moving average (ARMA) models to both wind speed and wind power output.

14. SUBJECT TERMS

wind energy; wind power; wind forecasting models; wind forecasting; autoregressive moving average (ARMA) models

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL